Federal Democratic Republic of Ethiopia Ministry of Health

Health Professionals' Competency Assessment and Licensure Directorate

Exam Development Manual

Message from the Director

As compared to the WHO standard of health professionals to population ratio for developing countries, Ethiopia has wide gap yet and the government has been working to increase the number of health professionals by increasing the intake of health science colleges and by opening new institutions.

While this helped to reduce the gap we have compared to the standard, the quality of the teaching and learning process hence, the service provided by the professionals became an issue for different stakeholders and the government as well. For this reason and since licensure exams are widely practiced internationally in different countries to insure the quality of health service, the federal ministry of health took the initiative to launch licensure exam initiative within human resource directorate since July 2015. Since then two round pilot tests given and reports discussed with different stakeholders.

The initiative currently led by a directorate established for this purpose, which is composed of three case teams and has been doing different activities to strengthen the system and capacitate the directorate to implement the exam in wider range with a much better organization and quality. As part of this, different documents were prepared to help facilitate the exam process and create transparency and sustainability of the program.

In the preparation of this document and other manuals, different examining institutions and countries experiences were reviewed, highly valued experts from different higher teaching institutions and Jhpiego-Ethiopia were involved for whom my deep appreciation goes for and for the high commitment and hardworking of the staffs at the federal ministry of health.

To this end, I would like to call for a collaborative work of all stakeholders in the health sector to this initiative, which indeed has high contribution to quality assurance of the health service delivery and yet cannot be effective unless all relevant bodies put their hands together for its implementation and continuity.

Dr. Ruth Nigatu

Health professionals' competency assessment and Licensure directorate, Director

Acknowledgments

Exam Development Manual

Abbreviations and Acronyms

CHD   Coronary Heart Disease

FDRE   Federal Democratic Republic of Ethiopia

FMoH   Federal Ministry of Health

HTN   Hypertension

KSA   Knowledge, Skill and Attitude

MCQ   Multiple Choice Questions

HPCALE   Health Professionals' Competency Assessment and Licensure Exam

HPCALD   Health Professionals' Competency Assessment and Licensure Directorate

OSCE   Objectively Structured Clinical Examination

SME   Subject Matter Expert

SPs   Standardized Patients

Table of Contents

Exam Development Manual

Exam Development Manual

List of Tables and Figures

Exam Development Manual

Section 1: Introduction

The World Health Organization (WHO) asserted that no matter how many individuals are educated and deployed; health professionals cannot transform population health unless they have the necessary competencies. Hence, the WHO recommended changes in regulations including certification and licensing of graduates. It is also known that many countries around the world including those in North America, Europe and Africa verify competence and fitness to practice of health professionals by administering standard qualification or licensure examinations.

The Government of Ethiopia recognizes that healthcare is one of the crucial components of basic social services with direct linkage to growth and development of the country as well as to the welfare of the society. The quality of health care delivered in a country has an immediate and long-term impact on the quality of life, livelihood, morbidity and mortality of its citizens and on the nation's economy and its national security. Accordingly, the Ministry of Health (MOH) is committed to reduce morbidity, mortality and disability and to improve the health status of Ethiopian people by providing and regulating health services. Health professionals' competency assessment is among the approaches of regulating health service; hence MOH designed health professional Competency assessment.

The sole purpose of competency assessment is to identify health professionals who possess the minimum basic knowledge and experience necessary to perform tasks on the job safely and competently, but not to select the "top" candidates. It is also to validate safe and effective medical practice, job-related skills and knowledge to provide an independent assessment and documentation of competency. Therefore, competency assessment is very different from academic or employment examination.

The purpose of exam development is to assure reliability and validity of exams by following important steps in the exam development process by applying an expertise made blueprint.

This manual is organized to describe the objective, development process, review, assembly system, standard-setting methods, and security-related issues to ensure the development of quality multiple choice questions and objective structured clinical examination (OSCE) stations.

Section 2: Objective

- ➢ The objective of this manual is to effectively guide the exam development process for standardized competency assessment of health professionals.

Section 3: Definition of Terms and Operational Definitions:

Blueprint: is a document which is used to represent the knowledge, skill, and attitude obtained from higher health institutions in a way that enables to measure the content, method of assessment and its weight.

Exam assembling: is a process of selecting and compiling items and stations according to the blueprint

Exam bank: a platform or system where exam item and stations are sorted, filtered and stored for future use

Exam developer: is a professional who prepares exam according to the exam blueprint.

Exam review: is a process of validating and reviewing exam.

Exam reviewer: is a professional who validates, identifies items and makes ready for final exam printing after the development of exam

International candidates: Are those candidates who have undertaken their educational training abroad and want to practice in Ethiopia

Multiple Choice Questions (MCQ): is one of a written assessment tool which enables the examinee to choose the single best answer from the list of options.

Objective Structured Clinical Examination (OSCE): is a tool for assessment of skills and attitude obtained from higher health institutions

Standardized patients (SPs): is someone who has been trained to portray, in a consistent, standardized manner, a patient in a medical situation.

Section 4: Roles and Responsibilities

4.1. Health Professionals' Competency Assessment and Licensure Directorate

- Plans, directs and implements competency assessment and licensure processes. In this case, the directorate controls the overall activities of the exam development case team.

4.2. Exam Development Case Team

- Plan and coordinate Task analysis process
- Plan, facilitate and execute exam blueprint preparation
- Design, Schedule and coordinate item development and construction of OSCE stations
- Plan and coordinate the exam review process
- Facilitate exam field testing process
- Guide and monitor the psychometric analysis,
- Collects large and easily accessible exam questions in a secure way.
- Compile and assemble a representative sample of the predetermined number of exam items based on the blueprint
- Keeps exam security in all steps of the exam development process
- Ensures efficient and effective workflow
- Carries out activities with other units, teams, and other offices in close cooperation
- Communicate the necessary information with other concerned bodies
- Prepare report and submit it to its immediate body
- Create, update documentation system

4.3 Exam Developer

- Constructs items or clinical case scenario based on the blueprint
- Maintain a high level of ethical principles during developing the exam
- Secure all information appropriately

4.4 Exam Reviewer

- Validate the developed exam items for content relevance, technical accuracy; clarity and sensitivity
- The editorial reviewer checks grammar, punctuation and spelling errors
- The reviewer will also verify exam fairness to all examinees and absence of culturally and religiously sensitive material to any particular group
- Maintain a high level of ethical principles during developing the exam
- Secure all information appropriately

4.5 Exam oversight committee

- The exam oversight committee will oversee the process of selection and training of item developers, reviewing and validating blue print and exam material, assembling of exam for each cadre and administration of exam and result dissemination

Section 5: Preparation Phase of Exam Development

Preparation for Exam Development process is targeted to select and train individuals who will be involved in the exam development and review activities. Based on the intended activity, selection criteria and the components of training provided may be different.

5.1. Selecting and Training of Exam Developers

The goal for selecting exam developers is to form a team of subject matter experts i.e. health professions education experts and psychometric experts (from different educational institutions and health facilities across the country) to develop high-quality MCQ items and OSCE stations.

To assure fairness, there shall be balanced institutional representation. Training of exam developer and reviewers should be based on a training manual that the directorate will prepare in the future.

5.1.1. Selection Criteria

Specific criteria for being exam developers

- Subject matter expert from different educational institutions and health facilities across the country
- Currently working in higher institutions as an instructor and having a Master's degree and above in the relevant field of study with a minimum of four years of experience as an instructor. For those disciplines instructed by BSc degree instructors; exam developers who hold first degree with instructing experience of four years can be selected
- Specialization in medical fields with a minimum of four years of experience
- Trained as an exam developer is preferable
- Can bring a support letter from his/her institution

5.1.2. Training of Exam Developers

After the selection of exam developers, standardized training on assessment principles, developing (designing) and reviewing high-quality MCQ items and OSCE stations shall be provided.

Training for MCQ item developers

- General information, specific details, and components specific characteristics about MCQ exam development
- Principles of evidence-based exam development techniques

For OSCE station developers

- General information, the role of assessors, specific characteristics of OSCE and principles about the station development

5.2. Selecting and Training of Exam Reviewers

5.2.1. Selection Criteria of exam reviewer

The reviewer:

- Should have a minimum of four years' experience after completion of a Master degree or specialty certificate in a specific field of study
- Should currently be delivering a course in the subject area (closely-related Course)
- Must bring a support letter from his/her working organization

NB: Before conducting an exam review, orientation shall be given for exam reviewers

Section 6: Exam Development Process

```
        ┌─────────────────────┐
        │     Conducting      │
        │   Task Analysis     │
        └─────────────────────┘
                  ↓
        ┌─────────────────────┐
        │     Blueprint       │
        │    Preparation      │
        └─────────────────────┘
                  ↓
        ┌─────────────────────┐
        │ Constructing Items  │
        │     & Stations      │
        └─────────────────────┘
                  ↓
        ┌─────────────────────┐
        │    Exam Revision    │
        └─────────────────────┘
                  ↓
        ┌─────────────────────┐        ┌──────────────┐
        │    Field Testing    │ ←───── │              │
        └─────────────────────┘        │ Modification │
                  ↓                     │              │
        ┌─────────────────────┐        └──────────────┘
        │    Psychometric     │        ┌──────────────┐
        │      Analysis       │        │     Not      │
        └─────────────────────┘        │   Selected   │
                  ↓                     └──────────────┘
        ┌─────────────────────┐        ┌──────────────┐
        │     Exam Review     │        │   Selected   │
        └─────────────────────┘        └──────────────┘
                                              ↓
        ┌─────────────────────┐
        │     Exam Banking    │
        └─────────────────────┘
                  ↓
        ┌─────────────────────┐
        │   Exam Assembling   │
        └─────────────────────┘
```
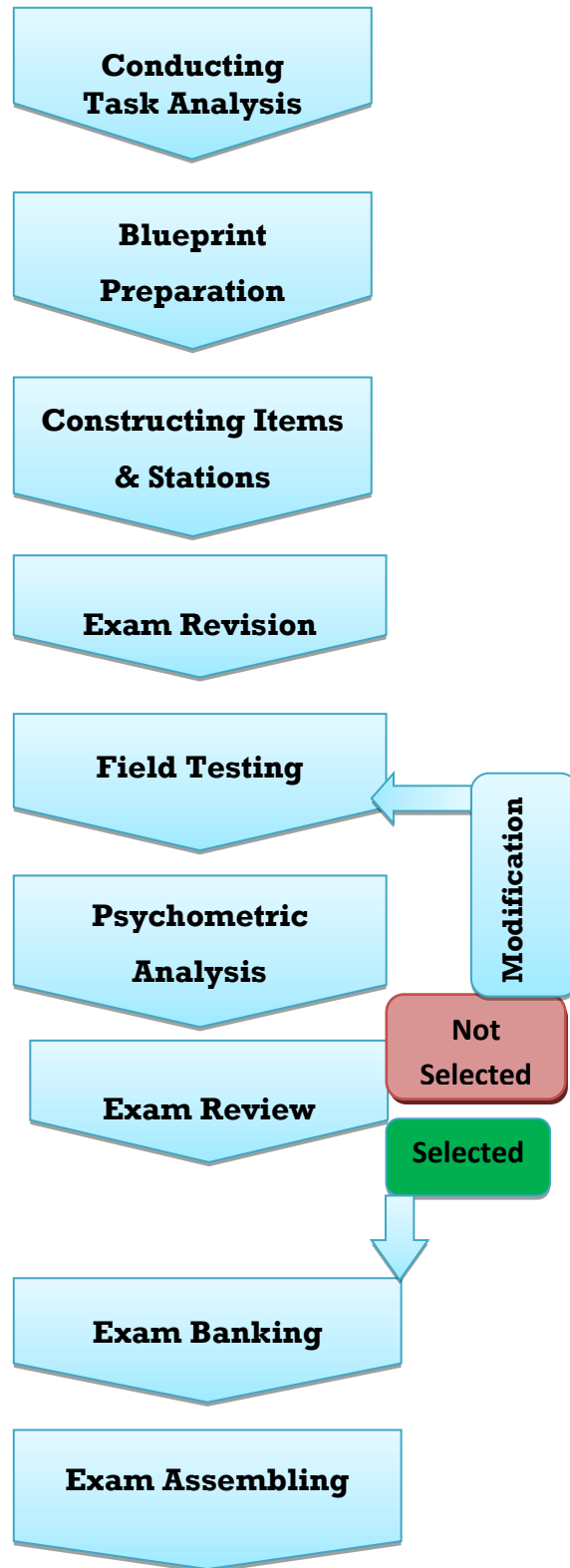
Figure1: Flow chart depicting the overall exam development process of HPCALE

6.1 Conducting Task Analysis

The purpose of task analysis in this manual is to systematically document the tasks that health care workers perform on-the-job; rather than relying on obsolete curricula or out-of-date national documents.

- HPCALD in collaboration with respective professional associations shall conduct the task analysis every 5 years but may do it more often to keep pace within change in health care services
- Task analysis should target those health care cadres who currently practice in their specific field of study that reflect current curricula and training

Task analysis involves two phases:

- Developing task list/task inventory through review of documents, scope of practice, curriculum, work place practices and expert validation
- Administering a survey with a representative sample of entry-level professionals

6.2 Exam Blueprint

- Two by two (content by a process) matrix is preferable
- It informs about the exam contents
- Communicates values of knowledge, skill, and attitude that are needed for effective practice and minimal competency
- Provide directions for individuals involved in exam construction (e.g. item/ case writers and station developers)

Components of Blueprint:

The exam blueprint matrix
1. The domains of competencies for the particular health profession
2. Core health problems or issues and professional tasks to be assessed
3. Relative emphasis/Weight of the content
4. Components of the task to be assessed such as learning domains (knowledge, skill, and attitude)
5. Exam format: MCQ and OSCE are preferable

Key steps in constructing a blueprint

The blueprint is prepared by plotting the program contents against the domains of the clinical competencies desired to be assessed. Blueprints are often written as a grid, usually content-by-process matrix *(Annex-1)*.

- Analyze and select competencies to be assessed
- Determine the basic function of the competency
- Identify the task (aspects of work) required to perform
- Identify the underpinning knowledge and skills required to perform the task
- Identify methods to collect evidence
- Identify the instrument (question or activities) test format
- Weight the content and decide the relative emphasis on percentage

6.3. Developing Exam

The health professionals' competency assessment and licensure exam includes the following components:

- Context and conditions of assessment
- Tasks to be administered to the candidates
- Outline of the evidence to be gathered from the candidates
- Evidence generation used to judge the level of performance

Instruction for Exam Developers

1. Stick to the manual and blueprint for constructing quality MCQ items and stations
2. Check for technically flawed items and tricky questions
3. Consider reviewer's feedback as needed
4. Make sure that an item has fulfilled all the technical and structural requirements before submission

6.3.1. Developing Multiple-Choice Questions:

> ➤ A single response (one best answer) type of MCQ item with four options should be used.

Components:

- Stem

- Lead-in

- Options/Alternatives, suggested solutions/:

  - **Key**: This is the correct choice for the item.

  - **Distracters**: These are the incorrect choices for the item

Sample Items

Basic Science Recall Item Stem: assesses recall of an isolated fact

- What area is supplied with blood by the posterior inferior cerebellar artery?

Basic Science Application of Knowledge Item Stem: assessing the application of knowledge.

- A 62-year-old man develops left-sided limb ataxia, Horner's syndrome, Nystagmus, and loss of appreciation of facial pain and temperature sensations. What is the most likely artery to be occluded?

Basic Characteristics of MCQ item

- Each item should focus on an important concept, typically a common or potentially catastrophic clinical problem

- Each item should assess the application of knowledge, not recall of an isolated fact

- The stem of the item must pose a clear question, and it should be possible to arrive at an answer with the options covered

- All distracters (incorrect options) should be homogeneous

- Avoid technical item flaws

- Make sure the item can be answered without looking at the options ("cover the option rule") OR that the options are 100% true or false

- The stems should be long and the options could be short

- Avoid superfluous information

- Avoid "tricky" and overly complex items

- Write options that are grammatically consistent and logically compatible with the stem

- List them in logical or alphabetical order

- Write distracters that are plausible and the same relative length as the answer

- Avoid using absolutes such as always, never, and all in the options; also avoid using vague terms such as usually and frequently

- Avoid negatively phrased items (e.g., those with except or not in the lead-in). If you must use a negative stem, use only short (preferably single word) options

6.3.2. Developing OSCE Stations
General Considerations of OSCE Stations Development

The OSCE stations shall have the following properties

- Components are assessed in a planned or structured way with attention being paid to the objectivity of the examination

- The station should be designed and developed by expertise which is assigned by OSCE team

- The development process of OSCE should be revised by exam reviewers

- Pilot testing of the station should be done before the stations are stored in the station bank.

- Each station should include a brief of the station profile like:
  - ➢ The objective of the station, station types, station date and duration, instructions, station resources, and opening statement etc.

- Each station is designed to assess a range of skills (practical skills, communication skill, data interpretation, and clinical decision-making skills) drawn from the elements of competence

- If necessary, an oral question may be applied and it should be well designed, more objective, scored in the separate sheet, and weight calculated accordingly

- A station may test one or a combination of different elements of competence

- The number of stations per exam should be 8 to 12 each lasting for 5 to 10 minutes

- The sequence of stations should be maintained in a way that stations will not affect each other

- All candidates pass through the same stations to maintain standardization

- When standardized/ simulated patients are used, detailed scripts are provided to ensure that the information that they give is the same to all candidates, including the emotions that the patient should use during the consultation
- There should be an instruction/ orientation package for examinee, examiners and standardized/ simulated patients
- There should be an organized team for designing and implementation of OSCE stations

6.3.2.1. Steps for OSCE station development

Step1. Developing a Blueprint for an OSCE

Step2. Converting the blueprint into a test map

Step3. Plan the details for each station

- Each OSCE station will have the following components

    1. **Station Profile**
- This profile is used for the documentation of the features of a station in the item bank. It is used as a title for the station. It clearly indicates an estimate of the time needed for the task, the domain, and competency tested and the logistics needed for the station *(Annex-2)*

    2. **Station Name**
- The station name should be given in clear and understandable way in line with the blueprint

    3. **Station Number**
- The station number should be assigned in a consecutive numbering

    4. **Station date and duration**
- The date of examination and duration of the task should be clearly stated on each station profile

    5. **Station types**
- The document should specify the type of the station for each task. Whether it is observable, technology enhanced, manned or unmanned

6. **Objective of the station**

- It is a statement that shows the task expected from the candidate and should demonstrate clearly

7. **Opening statement (Scenario)**

- The OSCE Scenario should be prepared based on common or critical patient presentations which the candidates will encounter as a practitioner/entry level health professional
  - ✓ It should provide relevant patient information to guide the examinee to perform the required task
  - ✓ The scenario should be written along with clear instruction
  - ✓ It should be written clearly (better to use patient language). (See the following example)

Example

A 54 years old man presented to ------ hospital OPD, complaining of chest pain of 1-month duration.

Perform focused chest examination

You have 10 minutes for this station.

---

8. **Station resources**

- The resources needed for each task should be clearly indicated

9. **Instructions**

- Instructions for examiners, examinees, and SPs should be clearly written and communicated timely

<u>**Instructions to the examinee**</u>

- It should be complete, stating clearly and precisely the setting, time allotted and the task that the examinee should perform
- Should be printed and laminated separately and given to the examinee or posted in a wall that can be easily read and re-read at a glance

Instructions for Examiners

Before the exam

- Should be briefed about the station profile before engaging them into the exam
- Should be instructed about the instructions given to the examinee and remind students to read the instructions if they appear to have forgotten what they should do
- Should be familiar with the checklists

During the exam

- Greet students in a similar way and don't coach students through facial expression and other non-verbal communication
- Intervene only for safety reasons
- Stay at the assessment area during the entire duration of the assessment activity and observe and record the students' performance on the checklist

After the exam

- Should complete the mark sheet and prepare to receive the next candidate,
- Record and report the result timely and provide feedback on the whole process of the OSCE to the Directorate using a score summary sheet *(Annex-3)*.

Instructions for the standardized and or simulated patients

- Should be brief, complete and clear; so that the standardized patients (SPs) portray the scenario consistently for all examinees.

Scoring and marking format

- For scoring and marking purpose, well designed and tested checklist rubrics should be used
- The checklist must have details about the expected items of the task to be performed
- Each item should be able to objectively assess candidate's performances and should have elaboration parameter *(Annex -4)*
- The number of items per checklist should not be more than 15 and not less than 5
- The Scoring checklist will have 3 columns rating scale; however, the OSCE preparing team can change it if necessary. (*Annex 5 and 6* of OSCE scoring sheets)

Items regarding generic skills which are expected to be in most or all stations should not be given a high or less mark as this will affect the overall mark of the exam.

Quality of OSCE Stations

- It should be evaluated using the station quality assurance checklist (*Annex-7*)

Section 7: The Exam Review Process

- ➢ The main purpose of the exam review process is for evaluating content relevance, technical accuracy, clarity and sensitivity related to culture and religion(***Annex-8 and Annex-9***)

Section 8: Conducting Field Testing, Psychometric Analysis and Review

8.1. Exam Field Testing or Pilot Testing

MCQ field testing and OSCE piloting will be conducted in randomly selected higher education institutions among graduating students.

Embedded Field Testing or Pilot testing

- The items are given in an actual test.
- Students are told that some of the items in their test will not count but are not told which items those are, so students try their best on all items.
- When field-tested items are embedded in the test, the invigilators shall collect information about how students will do on these items.
- Embedding field-test items to tests will either add additional items or replace actual test items to make room for the field-test items
- It identifies resource related gaps for actual testing

8.2 Psychometric Analysis
8.2.1. Item Analysis Statistics

- ➢ Item Difficulty Index and distracter analysis shall be employed to improve the items

8.2.2. Post OSCE analysis

Cronbach's alpha, G-theory, and Coefficient of R^2 psychometric analysis should be calculated and based on the result, station improvement suggestions have to be provided.

8.3 Post-pilot testing Exam Review

The analyses performed at the time of pilot testing are used to inform item reviewers in the following ways:

- The percentage of candidates responding correctly to an item (or the mean item score) informs test developers as to whether the item is more or less difficult for the students than intended.
- The percentage of candidates selecting each multiple-choice item response option informs test developers as to whether each potential distracter is sufficiently attractive to students.
- The relative number of candidates in each of the score categories of constructed-response items helps test developers to evaluate the scoring rubrics and definitions of the categories.
- It will help to adjust whether the duration of each OSCE station or the MCQ is sufficient or not

Section 9: Exam bank

9.1. Item Banking

This step is done after test items undergo a rigorous review and editorial process as well as field testing. Exams put in a bank are believed to accurately measure the knowledge, skill, and attitude of the candidate.

- Selection of items can be made through an automated item bank by randomly selecting items to meet pre-specified parameters.
- There are software programs that manage to file, sorting, storing, retrieval, statistical analysis and updating of items.
- As new items are added, others will have completed their service and will be ready to retire.
- The shelf-life of each item shall be ten years with slight modifications if necessary, after which it will be removed from the item bank.
- Items can be sorted and filtered to enable easy review by content experts and psychometric staff.
- Categorize items according to the content outline or blueprint

Information to be included during Item banking (at minimum)

- Unique item identifier
- Objective/learning outcome number from test blueprint
- Domain, (e.g. Dermatology)
- Category
- Subcategories
- Cases of an item
- Distracters/alternatives
- Testing dates-that an item was used and when it is scheduled for next use. The sequence number of item
- Number of candidates who attempted the item
- Classical statistics (Discrimination index and p-value/item difficulty)
- Average time to answer the item

- Author of item

- Item status-coded whether it is new or old

- Reference for answer verification

- Cognitive level of items across Bloom's Taxonomy

- Equivalent items (i.e. similar items that should not appear on the same form)

- Graphic link, if a graphic is part of an item

- Comments-of the psychometrician

- The frequency of items uses within five or ten years

Criteria for Item Banking:

➢ *Discrimination: Questions that fall into the Good category (greater than 0.3)*

➢ *Difficulty: Questions that fall into the Medium category (between 30% and 80%)*

➢ *Questions in the Poor category of a discriminatory level and questions in the Easy or Hard categories are recommended for further review.*

9.2. Station Banking

➢ Follows similar basic principles of MCQ banking mentioned in this guideline above.

The following information is to label each of the OSCE stations for banking

- Unique station identifier

- Objective/learning outcome number from test blueprint

- Station type (OSCE, Interpretation)

- Checklists / Rubrics

- Elaboration parameters

- Opening statement

- Examination dates that station was used and when it is scheduled for next use.

- The sequence number of station

- Number of candidates who attempted the OSCE station

- Classical statistics (Cronbach's alpha and Discrimination index)

- Average time to perform the station

- Author of station

- Station status-coded whether it is "new" or "old"

- Station resources (candidate, examiner's, simulated patient and non-standardized patients instructors, equipment list, special requirements e.g. restocking technician and also a scoring form

- Reference for verification

- Equivalent items (i.e. similar items that should not appear on the same form)

- Graphic link, if the graphic is part of an item

- Comments of the psychometrician

Criteria for Station Banking:

- A Cronbach's alpha of 0.7 to 0.8 should be taken as an acceptable level of reliability for the health professionals' competency assessment and licensure exam. The examination board should apply Cronbach's alpha that allows the detection of the OSCE stations which are main sources of error, by removing one station at a time from the analysis and looking at the reliability of the remainder.

- Multi-facet ANOVA analysis should be conducted to obtain the estimates of the components of variance for each of the following facets: stations, sites, test versions, and all their interactions. G-coefficient of 0.7 to 0.9 should be taken as an acceptable level of reliability for the health professionals' competency assessment and licensure exam.

- A good correlation ($R^2 > 0.5$) will indicate a reasonable relationship between checklist scores and global grades. The existence of low R^2 values at certain stations and/or widespread marks for a given grade should prompt a review of the item checklist and station design.

- When low R^2 reported, quality improvement should be undertaken byre-writing of the station and checklist with plans for re-use of this station and subsequent analysis of performance within a future OSCE.

Section 10: Exam Assembly

10.1. Exam assembling/formatting criteria

Items for inclusion on a test can be selected manually, randomly drawn by the computer from all existing items in the bank, or drawn by the computer from pre-specified parameters.

➢ Both MCQ and OSCE exams should also be assembled according to the exam blueprint and get ready for exam administration.

➢ When an item is selected or assembled for test inclusion, any graphics, tables, or cases associated with it should automatically be included.

➢ After completion of assembling the items, they shall be coded/ arranged in different **four** exam booklets in order to reduce the chance of cheating before administration of each exam.

10.2. MCQ Version Development

Exam versions are prepared by the exam assemblers using the computerized system and selection of items from exam bank. The exam format can also be structured by the computer software.

Item modeling and equating are the methods used for the preparation of equivalent exam version for a similar set of the examinee.

Item Modeling Process:

Item modeling is more successful with MCQs that have longer stems, especially clinical vignettes.

1. Select a source item- a well-written MCQ on a topic for which you want additional items
2. Highlight the specific terms in the stem that are important clinical content, (e.g., clinical setting, medical history, presenting complaint(s) etc.
3. Identify the correct (keyed) response and the content category to which it belongs
4. Review the available wrong options (distracters), and discard any that are inconsistent or flawed
5. For each clinically important term in the stem, list several significant alternatives

6. Prepare complete specifications for each new item. Identify the content of the new stem by labeling one clinically reasonable combination of the alternatives. Then, for each new stem, identify or provide a keyed response and stored to the bank

10.3. Exam Equating

- Exam equating is used to adjust candidates not to be advantaged by being assigned easier forms and disadvantaged by being assigned more difficult forms
- It is the statistical process used to adjust scores on test forms so that scores can be used interchangeably
- Equating adjusts for differences in difficulty among forms that are built to be similar in difficulty and content
- Generally, the following steps are suggested:
  - ✓ Choose a data collection design: single group randomly equated or non-equivalent
  - ✓ Get the parametric values such as difficulty index, discrimination index based on the classical test theory or item response theory.
  - ✓ Common item is selected from two tests with the same contents and same format.
  - ✓ The equivalent constant is calculated by a scale transformation

Section 11: Standard Setting

Standard setting is used to categorize candidates who score at or above the cut score and those who fail to score that. It should involve policymakers, test developers and content specialists.

Modified Angoff (1971) model is one of the most widely used standard settings that employ a test-centered approach for both written and skill licensure exams. Modified Angoff method is a judgmental approach in which a group of expert judges makes estimates about how borderline candidates would perform in the examination, i.e. the proportion of borderline examinees who will answer an item correctly.

11.1 Standard Setting for MCQ
Major Steps to be followed

Step 1: Select the raters:

- Select at least (7-15) Subject Matter Expert (SME) raters using the criteria for exam developers
- Assemble a diverse group of SMEs based on (e.g., gender, age, and educational background etc)

Step 2: Take the test:
- Have the raters take the test using the current cut score, if one has been established
- Obtain feedback from raters on objectives, wording, and design of test items
- If items need to be revised, do so before the rating process begins

Step 3: Rate the Items

Prior to beginning the actual rating process, conduct an orientation:
- Provide the definition of a "borderline candidate."
- Provide instructions on how to rate the test items
- Explain the rating process. Estimate the number of borderline candidate out of 100 who would answer the question correctly (Table 1)

Table 1: Sample test item rating form

TEST ITEM RATING			
COURSE NAME:		EXAM NAME	
RATER NAME:		Date:	
Instructions: Review each test item. Determine the probability that borderline candidates would answer the item correctly. Do not rate higher than 95% nor lower than 25%.			
TEST ITEM	Percentage (%) correct	MCQ ITEM	Percentage (%)correct
1		6	
2		7	
3		8	
4		9	
5		10	

Rating Process

- The raters will be given the test items along with the test item rating format above

- Raters are not provided with the answer key

- Separate the raters and have them provide estimates for each test item. Allow approximately two (2) hours for a 100-item test

- Reconvene raters and proceed to the next step

Step 4: Review the Ratings

- Collect the raters test item rating forms and enter the results Expert Ratings Spreadsheet (Table 2)

- Tabulate the average percentage correct for each test item by adding the raters' percentages and dividing by the number of experts

- Determine the standard deviation

- For any test item whose standard deviation exceeds 10, raters should discuss the reasons for variations in the estimates

- Finally, the cut score should get approval from the health professionals' competency assessment and Licensure Examination Committee

Test Items	Expert 1 Name	Expert 2 Name	Expert 3 Name	Expert 4 Name	Expert 5 Name	Average Percentage Correct	Standard Deviation
1							
2							
3							
4							
Average						Cut score	

Table 2: Sample Expert Ratings Spreadsheet

11.2 Standard Setting for OSCE

Major Steps to be followed:

Step 1: General orientation

- Scoring methods and any other information from which the panelists may benefit
- Where the faculty is not familiar with the OSCE stations, a mini OSCE is set as part of the standard setting orientation procedure
  - Faculty may play the role of examinees, or examiners, by observing each other
  - This is done to avoid overestimation or underestimation of the station difficulty. Their own level of performance as experts may serve as a `ceiling' effect for the standard-setting ratings.

Step 2: Orientation to a station

Station developers/ facilitators present the stations to panelists

- A full description of the stations
- Videotapes of one low performer and one high performer for the practice stations
- The actual skill score will be presented to the panelists following the completion of each video performance. The low vs. the high performers assist the panelist to form a range of possible performance profiles for the practice station

Step 3: Characteristics of borderline candidates

- Panelists should write independently the characteristics of the borderline candidate per skill component
- The panelists' statements are then posted and the facilitator discusses with the panelists each statement
- Arguments and disagreements are clarified and the group reaches a consensus as to what would be an appropriate list of borderline characteristics per skill component

Step 4: Panelists provide ratings

- Panelists are asked to make judgments as to how many items should be answered correctly by the borderline candidate in order to pass the stations
  - Note: in this manual, it is recommended that the number of items will constitute the ratings and not percentages of items.
- The facilitator should present all panelists' ratings on the board by assigning a number to each panelist.
- The panelists discuss their ratings
- The facilitator encourages panelists with the highest and lowest ratings to react to their judgments.
- The facilitator will average the ratings of the panelists to produce a cut-off raw score for the station.

Provision of actual performance data

- The distribution outlines the cumulative numbers and percentages of students who got one item correct, two items correct, three items correct etc.
- The facilitator should indicate the percentages of students who might fail the stations if the panelist average ratings are applied to the distribution as a cut-off score.
- A discussion should revolve around the issue of the `consequential data', i.e. percentage failure. Are panelists surprised? Should they expect a lower or a higher number of failures? Does the percentage failure seem to coincide with their experience? This provides another source of information, which helps panelists adjust their ratings when they are asked to attempt a second rating on the form
- Panelists provide their second ratings, which are posted on the board by the facilitator.

- A final cut-off score is calculated by averaging all the ratings

General consideration

Particularly, in the standard-setting process of performance assessment; to maximize the outcome of the standard-setting procedures it is recommended to have: -

- A large group of panelists (i.e. 18) into three groups of six each.
- The groups will set standards on different stations but one or two stations will be rated by all. This is done to examine the consistency of ratings among panelist groups

Station	Rater 1		Rater 2		Rater 3		Rater 4		Rater 5		Rater 6		Rater 7		Rater 8		Rater Average	an average of the station scores (student)
	1st Rating	2nd Rating	1st Rating	2nd Rating	1st Rating	2nd Rating	1st Rating	2nd Rating	1st Rating	2nd Rating	1st Rating	2nd Rating	1st Rating	2nd Rating	1st Rating	2nd Rating		
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
Passing score																		

Table 3: Sample standard setting spreadsheet for skill assessment

Section 12: Exam Security

- The exam bank will be kept in a secure location at the HPCALD and only the authorized personnel shall decide on who shall have access to the exam bank and when.

- Subject Matter Experts and other technical experts shall have restricted access as deemed necessary by the Directorate for the purpose of accomplishing their own respective tasks only.

- The HPCALD shall exercise Maximum Security in collaboration with National Educational Assessment and Examination Agency (NEAEA), Information Network Security Agency (INSA) and Addis Ababa University/Institute of Educational Research (AAU/IER) and other concerned authorities.

Exam security Spot Check

The standards of exam security will be monitored and upgraded at all steps of the competency assessment processes; during item development, field testing, review process, analysis, assembly, storage/banking. Assessing exam security, identifying problems and rectifying them are critical steps in ensuring the quality of exam security.

- During exam development, there needs to be a security camera in the exam development rooms.

- The exam will also be stored in a secure computer where the passwords are handled only by authorized staff and disabled from any of internet connection.

- To that end, regular spot check visits and review processes will be conducted at all steps.

- All individuals who are involved in any of the above processes are expected to sign an Exam Security Agreement *(Annexe-10)* and abide by the rules of the HPCALD with regards to overall exam security. Items shall be reviewed and edited by exam reviewers from different institutions

- The HPCALD will prepare an ethical charter/security agreement form and members and other individuals involved should sign and commit to the cause and shall be held accountable for any security breach.

Section 13. References

- Scott A. Cunningham, Sarah M. Callahan, Jason K. Feld, and John Richard Bergan, Item Development and Assessment Construction Guidelines for Common Core State Standards, 2013 (SQA 2018)

- Bluestone J, Johnson P, Fullerton J, Carr C, Alderman J, and Bon Tempo J. 2013. Effective in-service training design and delivery: evidence from an integrative literature review. Human Resources for Health 11:51. doi:10.1186/1478-4491-11-51.

- Leah Hart, PheleloMarole, Galina Stolarsky, (2014) Peter Johnson and Elizabeth Wagstaffe

- Linking Licensure Examinations to Nursing Practice through Task Analysis. A Case for BotswanaJhpiego Brown's Wharf 1615 Thames Street Baltimore, Maryland 21231-3492, USA www.jhpiego.org

- Oshio S, Johnson P, and Fullerton J. 2002. The 1999–2000 task analysis of American nurse-midwifery/midwifery practice. Journal of Midwifery & Women's Health 47(1): 35–41. doi:101.1016/S1526-9523(01)00218-5.

- AMCB. 2012 Task Analysis Report: A Report of Midwifery Practice. Linthicum, MD: American Midwifery Certification Board; 2012.

- The Federal Democratic Republic of Ethiopia, Ministry of Health, 2015. Draft Implementation Guideline for health professionals' competency assessment and Licensure Examination of Health Professionals. FDRE MoH.

- Airasian, P. (1994) Classroom Assessment, Second Edition, NY: McGraw-Hill.

- Brame, C., (2013) Writing good multiple choice test questions.

- Brown, F. G. (1983). Principles of educational and psychological testings (3rd Ed.). New York: Holt, Rinehart, and Winston

- Cangelosi, J. (1990) Designing Tests for Evaluating Student Achievement. NY: Addison Wellesley.

- Cassin, W. E. (1987). Improving essay tests. Idea Paper, No. 17. Manhattan, KS: Center for Faculty Evaluation and Stevens,

- Greenland, N (1993) How to make achievement tests and assessments, 5th edition, NY: Allen and Bacon.

- Implementation Guideline for health professionals' competency assessment and Licensure Examination of Health Professionals,2015

- Wickens, T. D., Multiday contingency tables for the social sciences. Hillsdale, NJ: Lawrence Erlbaum Associates (1989) p. 238.

- Williams, J. R., Tiller, M. G., Herring, H. C. III, & Steiner, J. H. (1988). A framework for the development of accounting education research. Sarasota, Florida: American Accounting Association.

- Burton, S. et al., 1991. How to prepare better multiple-choice test items: Guidelines for university faculty. *Brigham Young University Testing Services and The Department of Instructional Science*, p.36. Available at: http://testing.byu.edu/info/handbooks/better items.pdf\nhttp://www2.swccd.edu/~mseteachingresources/assets/multiple choice/multiplechoiceguidelines.pdf.

- Callahan, S.M. & Ph, D., 2013. Item Development and Assessment Construction Guidelines for Common Core State Standards by.

- Cohen, A.S., Handbook on Test Development : Helpful Tips for Creating Reliable and Valid Classroom Tests and James A. Wollack Testing & Evaluation Services University of Wisconsin-Madison 1. Terminology Used in Testing. *Test*.

- Medical Council of Canada, 2013. Guidelines for the development of objective structured clinical examination (OSCE) cases, (November), pp.2–36.

- Mullis, I.V.S. & Martin, M.O., 2011. TIMMS 2011 Item-Writing Guidelines. , pp.1–28.

- Prince Georges Community College, Best Practices for Writing Multiple-Choice Questions Best Practices for Writing Multiple-Choice Questions. In pp. 1–3.

- Saudi Commission for Health Specialties, 2014. *Objective Structured Clinical Examination OSCE*,

- Shirwaikar, A., 2015. Objective structured clinical examination (OSCE) in pharmacy education- A trend. *Pharmacy Practice*, 13(4), pp.1–5.

- Isaac I. Bejar., Standard Setting: What Is It? Why Is It Important? Copyright © 2008 by Educational Testing Service. http://www.ets.org/research/contact.html

- Cizek, G. J., & Bunch, M. B. (2007). Standard setting: A guide to establishing and evaluating performance standards on tests. Thousand Oaks, CA: Sage.

- AMEE Guide No. 18. 2000. Miriam Friedman Ben-David: Standard setting in student assessment Medical Teacher, Vol. 22, No. 2

- Angoff, W.H. (1971). Scales, norms and equivalent scores. Educational Measurements.

Washington, DC: American Council on Education

- Alan Wheaton and Jim Parry, (2012). Using the Angoff Method to Set Cut Scores. Test Defensibility in the US Coast Guard - Using the Angoff Method to Set Cut Scores

- Kane MT. 2006. Validation. In: Brennan RL, editor. Educational Measurement. Westport, CT: American Council on Education and Praeger Publishers. pp 17–64

- Ben-David MF. 2000. AMEE Guide No. 18: Standard setting in student assessment. Med Teach 22:120–130.

- Epstein RM. 2007. Assessment in Medical Education. N Eng J Med 356:387–396.

- The Federal Democratic Republic of Ethiopia, Ministry of Health, November 2015. Implementation Guideline for health professionals' competency assessment and Licensure Examination of Health Professionals.

- Federal Ministry of health professionals' competency assessment and licensure examination for health professional directorate. OSCE Implementation guideline, 2017

- Hays R. 2008. Assessment in medical education: Roles for clinical teachers. Clin Teach 5:23–27.

- Hossam Hamdy 2006. Blueprinting for the assessment of health care professionals, Practical

- Assessment. Blackwell Publishing Ltd 2006. THE CLINICAL TEACHER 2006; 3: 175–179

- Kamran z. Khan. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective

- Mary E. McDonald, 2007. The Nurse Educator's Guide for Assessing Learning Outcomes.2nd edition.Jones and Bartlett Publishers.

- Saudi Objective Structured Clinical Examination (OSCE) MANUAL. Saudi Commission for Health Specialties, Riyadh – 2014, ISBN: 978-9960-9832-8-8

- Steven M Downing, 2003. Validity: on the meaningful interpretation of assessment data. Medical Education 37:830–837

- Sood R. 2001. Long Case Examination - Can it be improved. Indian AcadClin Med 2:251–255

Section 14 Annexes

14.1. Annex-1 Blueprint Example for Midwifery Profession (Obstetric II Module)

Competency Area	Basic Science	Assessment	Diagnosis	Management	Evaluation	Total
Abnormal Pregnancy	4.3	6.17	4.67	5.67	3.17	24
Medical Problems During Pregnancy	3.17	2.67	2.5	2.8	2	13
Abnormal Labor & Delivery	4.67	6.67	5.3	7.3	4	28
Obstetric Emergencies	2.3	4.5	4.17	4.83	2.17	18
Abnormal Purperium	2.5	3.67	4	4.3	2.5	17
Total						100

Module competency	Task per Competency	Learning outcome	Assessment method	Emphasis per LO	Emphasis per task
Manage abnormal pregnancy	**Basic Science**	• Describe the pathophysiology of abnormal pregnancy	MCQ	1.1	**4.3**
		• Explain the types of abnormal pregnancy	**Short Answer**	1	
		• Analyze the consequence of abnormal pregnancy	MCQ	1.1	
		• Identify sign symptoms of abnormal pregnancy	MCQ	1.1	
	Assessment	• Take an appropriate history of AP	OSCE	2	**6.17**
		• Perform proper physical examination AP	OSCE	2	
		• Send lab request accordingly	MCQ/ OSCE	1.1	
		• Order U/S	MCQ	1.07	
	Diagnosis	• Identify diagnostic methods for AP	MCQ	1.67	**4.67**
		• Differentiate different pregnancy abnormalities	MCQ	1.5	
		• Identify D.Dx for each pregnancy-related abnormalities	MCQ	1.5	
	Management	• Analyze different management principles of AP	MCQ		**5.67**
				1	
		• Describe medical and surgical management of AP	MCQ	0.75	

		• Manage abnormal pregnancy	MCQ/ OSCE	1.75	
		• Counsel the mother about possible measures on AP respectfully	OSCE	1	
		• Discuss prevention principles	MCQ	1.17	
	Evaluation	• Follow the patient condition after intervention	OSCE	1	3.17
		• Judge the patient's intervention pertaining to AP	MCQ	0.67	
		• Recommend the alternative interventions for a patient with AP	MCQ	0.5	
		• Decide on the patient's progress	MCQ	1	
Manage Medical problems of pregnancy	Basic Science	• Explain the pathophysiology Md. D/O during pregnancy	MCQ	1	3.17
		• Describe risk factors and prevention mechanisms related to Medical problems during pregnancy	MCQ	0.77	
		• Identify sign and symptoms of M. D/o	MCQ	0.9	
		• Analyze the consequences of M.D. of pregnancy	MCQ	0.5	
	Assessment	• Take full Hx	OSCE	0.92	2.67
		• Perform P/E	OSCE	0.82	
		• Approach a patient in a friendly manner during Hx and P/E	OSCE	0.5	
		• Order appropriate lab investigation	MCQ/OSCE	0.4	
	Diagnosis	• Identify m/d/o during pregnancy	MCQ	1	2.5
		• Identify the diagnostic methods for Medical problems during pregnancy	MCQ	1	
		• Interpret lab results	MCQ	0.5	
	Management	• Describe the management of Medical problems during pregnancy	MCQ	0.8	2.8
		• Provide appropriate care for pregnant woman with a medical problem in a	MCQ/OSCE	1.25	

		• respectful manner			
		• Provide HE for a pregnant woman about M.DO	**OSCE**	**0.75**	
	Evaluation	• Follow the patient condition after intervention	**OSCE**	**0.6**	**2**
		• Decide on alternative measures when primary intervention has failed	**MCQ/ OSCE**	**0.3**	
		• Evaluate the current condition of a patient after intervention	**MCQ/ OSCE**	**0.4**	
		• Link with other services if necessary	**MCQ/ OSCE**	**0.2**	
		• Counsel when to return about the patient follow up	**OSCE**	**0.5**	
Manage abnormal labor and delivery	**Basic Science**	• Explain the pathophysiology during abnormal labor and delivery	**MCQ/Essay**	**1.27**	**4.67**
		• Describe risk factors and prevention mechanisms of abnormal labor and delivery.	**MCQ? Essay**	**1.1**	
		• Identify sign and symptoms of abnormal labor and delivery.	**MCQ**	**1.1**	
		• Analyze the consequences of abnormal labor and delivery.	**MCQ/Essay**	**1.2**	
	Assessment	• Take full Hx	**OSCE**	**2**	**6.67**
		• Perform P/E	**OSCE**	**1.67**	
		• Approach a patient in a friendly manner during Hx and P/E	**OSCE**	**1**	
		• Order appropriate lab investigation	**OSCE**	**1**	
		• Order U/S	**OSCE**	**1**	
	Diagnosis	• Identify m/d/o during abnormal labor and delivery	**MCQ**	**2**	**5.3**
		• Identify the diagnostic methods for abnormal labor and delivery	**MCQ/ Essay**	**2**	
		• Interpret lab results	**MCQ/Essay**	**1.3**	
	Management	• Describe the management of abnormality during	**MCQ/Essay/S hort answer**	**2.6**	

					7.3
		• Provide appropriate care for women comes with an abnormality in labor in a respectful manner	**OSCE**	**2.7**	
		• Provide emotional support for a mother during abnormal labor and delivery	**OSCE**	**2**	
	Evaluation	• Follow the patient condition after intervention	**Direct Observation**	**1**	**4**
		• Decide on alternative measures when primary intervention has failed	**MCQ/Short Answer**	**0.75**	
		• Evaluate the current condition of a patient after intervention	**MCQ/OSCE**	**0.75**	
		• Link with other services if necessary	**OSCE**	**0.5**	
		• Counsel when to return about the patient follow up	**OSCE**	**1**	
Identify and manage obstetrics emergency	**Basic Science**	• Describe the pathophysiology of obstetric emergency	**MCQ/Essay**	**0.6**	**2.3**
		• Explain types of obstetric emergency	**Essay/Short Answer**	**0.4**	
		• Analyze the consequence of obstetric emergency	**Essay**	**0.8**	
		• Identify sign symptoms of an obstetric emergency	**MCQ**	**0.5**	
	Assessment	• Take full Hx	**OSCE**	**1.25**	**4.5**
		• Perform P/E	**OSCE**	**1.25**	
		• Approach a patient in a friendly manner during Hx and P/E	**OSCE**	**1**	
		• Order appropriate lab investigation	**MCQ/OSCE**	**0.5**	
		• Order U/S	**OSCE**	**0.5**	
	Diagnosis	• Identify obstetrics emergency	**MCQ**	**1.5**	**4.17**
		• Identify the diagnostic methods for obstetric emergency	**MCQ**	**1.5**	
		• Interpret lab results	**MCQ**	**1.17**	
	Management	• Describe the management of obstetric emergency	**MCQ**	**1.5**	

		• Provide appropriate care for patients with obstetrics emergency in a respectful manner	OSCE	1.83	4.83
		• Provide HE for a pregnant woman about obstetrics emergency	OSCE	1.5	
	Evaluation	• Follow the patient condition after intervention	Direct Observation	0.67	2.17
		• Decide on alternative measures when primary intervention has failed	MCQ/OSCE	0.4	
		• Evaluate the current condition of a patient after intervention	OSCE	0.4	
		• Link with other services if necessary	MCQ/ OSCE	0.3	
		• Counsel when to return about the patient follow up	OSCE	0.4	
Manage abnormal puerperium	Basic Science	• Describe the pathophysiology of abnormal puerperium	MCQ/Essay	0.75	2.5
		• Explain types of abnormal puerperium	Essay	0.5	
		• Analyze the consequence of abnormal puerperium	Essay	0.75	
		• Identify sign symptoms of abnormal puerperium	MCQ	0.5	
	Assessment	• Take full Hx	OSCE	0.9	3.67
		• Perform P/E	OSCE	0.87	
		• Approach a patient in a friendly manner during Hx and P/E	OSCE	0.7	
		• Order appropriate lab investigation	OSCE	0.7	
		• Order U/S	OSCE	0.5	
	Diagnosis	• Identify abnormal puerperium	MCQ/Short Answer	1.5	
		• Identify the diagnostic methods for abnormal puerperium	MCQ/Short Answer	1.5	4
		• Interpret lab results	Short Answer	1	
	Management	• Describe the management of abnormal puerperium	MCQ/Essay	1.4	4.3
		• Provide appropriate care for patients with abnormality during puerperium obstetrics emergency in a respectful	OSCE	1.49	

		manner			
		• Provide HE for a pregnant woman about abnormal puerperium	**OSCE**	**1.4**	
	Evaluation	• Follow the patient condition after intervention	**Direct Observation**	**0.57**	**2.5**
		• Decide on alternative measures when primary intervention has failed	**MCQ/ OSCE**	**0.4**	
		• Evaluate the current condition of a patient after intervention	**OSCE**	**0.4**	
		• Link with other services if necessary	**MCQ/OSCE**	**0.3**	
		• Counsel when to return about the patient follow up	**OSCE**	**0.5**	

14.2. Annex-2 Sample OSCE Station profile template

Station Name:		Station #
Station Type: **Observed** ☐ **Un observed** ☐	Date: Time Given:	
Station Domain:		
Station Objectives:		
Scenario:		
Domain/ Skills to be Performed		
Resources needed *(list all the necessary resources for this particular station: equipment, standardized patients)*		
Instruction for the Examiner		
Instruction for Standardized Patients (if any)		
Instruction for Students		

14.3. Annex-3 Score Summary Sheet

Station Name:	**Station #:**	
Student ID:	**Date:**	
OSCE Total Score (100%)		
Oral Examination Total Score (100%)		
Overall Score (100%)		

Station Pass Score (SPS):
Station Clear Pass Description:

Examiner Overall Judgment:

Alternative I SPS

Clear Fail ☐	**Borderline** ☐	**Clear Pass** ☐	**Superior** ☐
(less than pass score	(±1SD from SPS)	(+2SD from SPS)	(+3SD and above from
SPS)			

Alternative II: Description

Clear Fail ☐	**Borderline** ☐	**Clear Pass** ☐	**Superior** ☐

Examiner Name and Signature:_____

Students Signature: _____

14.4. Annex-4 Elaboration parameter for each activity in the checklist

S.Nº	Checklist Step/Task	Elaboration Parameter		
		2	1	0
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

14.5. Annex-5 Scoring and Marking Tool (Checklist)

Station Name: _____ Station #: _____

Student ID: _____

	Task/Activities/Items	Rate/Score		
		Not attempted at all	*Attempted but not satisfactorily*	*Satisfactory*
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
11.				
12.				
13.				
14.				
15.				
OSCE Total Score (100%)				

14.6. Annex-6 Oral Questions for OSCE

Question No.	Questions	Possible Answers with their respective Score		
1.				
2.				
3.				
4.				
Total Oral Examination Score (100%)				

Examiner Feedback *(Both for OSCE and Oral):*

_____

_____

_____

14.7. Annex 7: OSCE station Quality assurance checklist

Station Author/s:

Station Title:

No.	STATION COMPONENT:	Check Points	Yes	No
1.	Name	Does the station name given in clear and understandable way in line with the blueprint matrix?		
2.	Objective	Does the objective state clearly in a statement that shows of the task expected the candidate should demonstrate?		
3.	Opening statement (Scenario)	Does the scenario provide relevant patient information to guide the examinee to the required task with clear instruction?		
4.	Station Resources	Are the resources needed for each task clearly indicated?		
5.	Instructions	Are the instructions for examiners clearly written?		
		Are Instructions for examinees clearly written?		
		Are Instructions for SPs clearly written?		
6.	Scoring and marking format	Does the checklist have key items that can be assessed objectively		
		Are the number of items not more than 15 and less than 5		

14.8. Annex 8: Checklist for reviewing Multiple-Choice Items

Item Type	
Date Written	
Item Developer	
Reviewers	
Review Date	

Areas to be Reviewed	Activities	Check
Content	The item addresses the desired domain of knowledge and corresponds to the exam blueprint.	
	It measures knowledge or a skill component which is worthwhile and appropriate for the examinees who will be tested	
	There is a no better way to test what this item tests	
	The level of difficulty is appropriate for the examinees that will be tested.	
	The item is focused on current standards of practice	
	The item tests higher order thinking.	
	No grammatical, punctuation, and spelling errors.	
	The item is free from demeaning, offensive, or stereotypical content	
Issues Related to Test wise-ness	Grammatical cues - one or more distracters don't follow grammatically from the stem	
	Logical cues - a subset of the options are collectively exhaustive	
	Absolute terms - terms such as "always" or "never" are in some options	
	Long correct answer - the correct answer is longer, more specific, or more complete than other options	
	Word repeats - a word or phrase is included in the stem and in the correct answer	
	Convergence strategy - the correct answer includes the most elements in common with the other options	
Issues Related to Irrelevant Difficulty	Options are long, complicated, or double	
	Numeric data are not stated consistently	
	Terms in the options are vague (eg, "rarely," "usually")	
	Language in the options is not parallel	
	Options are in a non-logical order	
	"None of the above" is used as an option	
	Stems are tricky or unnecessarily complicated	
	The answer to an item is "hinged" to the answer of a related item	
Correct (Keyed) Answer	Placement of the correct answer has been varied	

Comments: _____

14.9. Annex-9 Checklist for reviewing OSCE

Station Type	
Date developed	
station Developer	
Reviewers	
Review Date	

Areas to be Reviewed	Activities	Check
OSCE station	The prepared station corresponds to the exam blueprint	
	Examinees sheet is prepared that includes the station profile, the opening statement (scenario or preface) and the candidate instructions	
	Examiner's sheet contains the station profile and the examiner's instructions accompanied by a scoring form (checklist or rating scales)	
	Tasks and instructions for the support team are prepared to describe the members of the team and their responsibilities	
	Instructions to the standardized patient are detailed enough	
	Station labels, instructions and directions are prepared	
	All necessary equipment to perform the procedure are available at the station and are functional	
The OSCE scenario	Prepares the mindset of the examinee and provides a semi-real-life situation	
	Stated in the patient's language rather than technical medical terminology	
	Provides sufficient information to guide the examinee to the required task (not too much, not too little)	
The OSCE checklist	The number of items on the checklist is proportional to the case and the time allotted	
	Each item represents only one concept	
	Items on the checklist are observable skills	
	Each checklist item begins with an action verb	
	The item focuses on current standards of practice	
	Contains an explicit marking system	

Comments:

_____
_____
_____.

14.10. Annex-10 Test Security Agreement

> ### *For team of exam development, exam review, standard setting or assembling of the exam for the HPCALE*

As a member of the team for the exam development, standard setting, exam review or assembling of the Health Professionals' Competency Assessment and licensure exam in Ethiopia, I, the undersigned, accept the responsibility for maintaining strict confidentiality of items, stations, materials and information related to the examination as detailed below:

I declare that no one from my family members, including my spouse, children, brothers, and sisters, is a candidate for the licensure examination in the year of this exam development.

I will not share any information about the examination with anybody including my families and friends except duly authorized persons (i.e., team members from HPCALD and AAU/IER), though any means or media including telephone, text message, pictures, internet, print materials, face to face discussion, etc.

I am aware that examination items under my control including, but not limited to scratch papers, draft copies, are to be kept in a secure location.

I am aware that I may be sued in accordance with the country's law if I violate these provisions.

I have read and understood the provisions of this security agreement. My signature below signifies that I do accept the terms and conditions of this security agreement.

Full Name: _____

Signature: _____

Date: _____/_____/_____